

**UNITED STATES DISTRICT COURT
DISTRICT OF NEW JERSEY**

IN RE: INSULIN PRICING LITIGATION

**Case 2:23-md-03080
MDL No. 3080**

**JUDGE BRIAN R. MARTINOTTI
JUDGE RUKHSANAH L. SINGH**

DECLARATION OF DOUGLAS FORREST

I, Douglas Forrest, state and declare as follows:

1. I am the Senior Vice President, eDiscovery Analytics & Strategy, at International Litigation Services (“ILS”), which is located in Irvine, California (www.ilsteam.com). I have been retained as a consultant for the Plaintiffs in this action. The facts stated in this declaration, except as otherwise explicitly noted, are within my own personal knowledge and, if called as a witness to testify, I could and would competently testify to the facts contained in this declaration.

2. I make this declaration in support of Plaintiffs’ Proposed ESI Order. In preparing this declaration I reviewed:

- a. Plaintiffs Proposed ESI Order (Ex. 1)¹;
- b. Defendants’ Proposed ESI Order (Ex. 2);
- c. *In Re: Allergan Biocell Textured Breast Implant Liability Litigation*, Case 2:19-md-02921, MDL No. 2921 (D.N.J. Sept. 4, 2020), ECF No. 194, Case Management Order No. 15 (Order Regarding Electronically Stored Information and Hard Copy Documents) (Ex. 4);
- d. *In Re: 3M Combat Arms Earplug Product Liability Litigation*, Case 3:19-md-02885, MDL No. 2885 (N.D. Fla. June 17, 2019), ECF No. 443

¹ All citations to “Ex. ___” refer to exhibits to the Declaration of David R. Buchanan dated March 22, 2024.

Pretrial Order No. 10, Order Governing Production of Documents and Electronically Stored Information (Ex. 11); and

- e. *In Re: Volkswagen “Clean Diesel” Marketing, Sales Practices, and Products Liability Litigation*, Case 3:15-md-02672, MDL No. 2672 (N.D. Cal. May 4, 2016), ECF No. 1482, Pretrial Order No. 18 (Stipulation and Order Governing the Production of Hard Copy Documents and Electronically Stored Information) (Ex. 12).

I. QUALIFICATIONS

3. I am a graduate of Stanford Law School, where I was a Note Editor of the Law Review. I was admitted to the bar in 1977 (I currently have retired status), and, after practicing law at Breed, Abbott & Morgan and Cravath, Swaine & Moore, I developed expertise in computer technology and software design, programming, and implementation, both generally and with respect to litigation support and e-discovery.

4. As an attorney at Cravath, I relied on *Aquarius*, the first large-scale implementation of computerized litigation support, which was implemented on the IBM antitrust cases.

5. As Director of Litigation Services at Legal Information Technology, Inc. (“LIT”), I was instrumental in introducing imaging, coding, and search technology for discovery to Am Law 200 law firms, and I pioneered the practice of integrating imaging with legacy search systems such as BRS.

6. As a systems architect, application designer, and programmer, I created case management, litigation support, and document repository systems and applications (including *WIDE*, and *LIT CaseWorks for Lotus Notes*, developed for LIT), SaaS (Software as a Service) knowledge management applications (including *LexisNexis Total Alerts* and *LexisNexis Clipper* developed for Ozmosys (<https://www.ozmosys.com/>)), and e-discovery and production operation systems for Doar (<https://www.doar.com/>). The software that I designed, programmed, and

implemented to produce *LexisNexis Total Alerts*, *LexisNexis Clipper*, *LexisNexis Legal Industry Monitor*, *Thompson Elite Daily Docket*, and *Institutional Investors' Mutual Fund Daily*, *Hedge Fund Daily* and *Compliance Daily*, traversed thousands of web sites and extracted thousands of URLs from those sites daily.

7. At ILS, I direct e-discovery analytics and strategy, providing technical advice, expertise, and drafting and other assistance to counsel.

8. I have advised, consulted, or acted as a declarant or affiant with respect to ESI in many cases, including:

- a. *In re: Uber Technologies, Inc. Passenger Sexual Assault Litigation*, MDL No. 3084 (N.D. Cal.);
- b. *In re: Social Media Adolescent Addiction/Personal Injury Products Liability Litigation*, MDL No. 3047 (N.D. Cal.);
- c. *In Re: StubHub Refund Litigation*, Case No. 4:20-md-02951-HSG (N.D. Cal.);
- d. *In Re: Meta Pixel Healthcare Litigation*, Case No. 3:22-cv-3580-WHO (N.D. Cal.);
- e. *In Re: Philips Recalled CPAP, Bi-Level Pap, and Mechanical Ventilator Products Litigation*, MDL No. 3014 (W.D. Pa.);
- f. *Nichols v. Noom Inc.*, No. 20-CV-3677 (LGS) (KHP), 2021 WL 948646, (S.D.N.Y.);
- g. *In re: Ethiopian Airlines Flight ET 302*, Lead Case: 1:19-cv-02170 (N.D. Ill.) (Boeing 737 Max Crashes);
- h. *In Re: 3M Combat Arms Earplug Product Liability Litigation*, MDL No. 2885 (N.D. Fla.) (“3M Combat Arms MDL”);
- i. *In Re: Volkswagen “Clean Diesel” Marketing, Sales Practices and Products Liability Litigation*, MDL No. 2672 (N.D. Cal.) (“VW Clean Diesel MDL”);
- j. *In Re: Intel Corp, CPU Marketing, Sales Practices and Products Liability Litigation*, MDL No. 2828 (D. Or.);
- k. *In Re: Chrysler-Dodge-Jeep EcoDiesel Marketing, Sales Practices and Products Liability Litigation*, MDL No. 2777 (N.D. Cal.);

- l. *Lafferty v. Alex Jones* (Conn. Super. Ct.);
 - m. *Soto v. Bushmaster Firearms International* (Conn. Super. Ct.) (Sandy Hook parents);
 - n. *In Re: Takata Airbags Product Liability Litigation*, MDL No. 2599 (S.D. Fla.);
 - o. *In Re: Testosterone Replacement Therapy Products Liability Litigation*, MDL No. 2545 (N.D. Ill.);
 - p. *In Re JCCP 4771, Zolof Birth Defect Cases*, (Cal. Super. Ct.);
 - q. *Da Silva Moore v. Publicis Groupe & MSL Group*, No. 11 Civ. 1279 (ALC)(AJP) (S.D.N.Y.) (seminal TAR case).
9. I have served as a speaker or panelist on many CLE webinars including *Balancing the Needs of Requesting and Producing Parties: Getting E-Discovery Right* (RAND Institute for Civil Justice Conference, October 3-4, 2023) (“RAND Conference”), *An Analysis of Today’s Mass Tort Landscape Agenda* (HarrisMartin MDL Conference, March 27, 2019), *Current Mass Torts from E-Discovery Through Exit Strategies - Navigating “Game-Changing” Dynamics* (HarrisMartin MDL Conference, November 26, 2018), *The Mass Tort Litigation Landscape - A Critical Analysis Agenda* (HarrisMartin MDL Conference, September 26, 2018), *The State of E-Discovery in 2018: Analysis & Review* (West LegalEdCenter, September 27, 2018), *Lessons Learned from Recent eDiscovery Disasters* (West LegalEdCenter, February 26, 2018), and *Top ESI Mistakes Made in Mass Tort Disputes* (West LegalEdCenter, September 14, 2017).
10. I was an advisor on and assisted in setting up the RAND Conference.
11. I am a member of the drafting team for forthcoming commentary of the Sedona Conference Working Group 1 (<https://thesedonaconference.org/wgs/wg1>) on *Conducting eDiscovery of Modern Communication and Collaboration* and was on the panel presenting the *Drafting Team Report: Discovery of Modern Communications and Collaboration Platforms* at the Sedona Conference Working Group 1 Annual Meeting in October 2023.

II. OPINIONS

The Deficiencies of Search Terms

12. The deficiencies of search terms in litigation were exposed in the 1985 publication of the landmark paper by Blair and Maron² that found that only 20% of relevant documents included search terms selected by experienced counsel (who thought that the search terms would have retrieved 75% of relevant documents).

13. Later studies at the 2007 and 2008 TREC Legal Track³ confirmed this dismal state of affairs, finding estimated recall⁴ rates for search terms of 22%, i.e., 78% of relevant documents missed, and 24%⁵, missing 76% of relevant documents. Perhaps worse, with respect to highly relevant documents, the result of the application of a “consensus set” of Boolean search terms in the 2008 Legal Track implied “that, on average per topic, 58% of the ‘highly relevant’

² David C. Blair & M. E. Maron, An Evaluation of Retrieval Effectiveness for a Full-Text Document-Retrieval System, 28 COMM’NS ACM 289, 290 (1985).

³ The Text Retrieval Conference (TREC), co-sponsored by the NIST Information Technology Laboratory’s (ITL) Retrieval Group was established in 1992. Celebrating 25 Years of TREC, available at <https://trec.nist.gov/celebration/25thcelebration.html>. The Legal Track at the Text REtrieval Conference (TREC) was established “to assess the ability of information retrieval techniques to meet the needs of the legal profession for tools and methods capable of helping with the retrieval of electronic business records, principally for use as evidence in civil litigation.” TREC Legal Track, available at <https://trec-legal.umiacs.umd.edu/>

⁴ In the discovery context, recall is the percentage of responsive documents identified by a classification process. When applied to search terms, recall is the percentage of responsive documents that contain a search term. With TAR, recall is the percentage of responsive documents classified as responsive by the TAR process.

⁵ Stephen Tomlinson, et al., Overview of the 2007 TREC Legal Track (April 30, 2008), (22% recall), available at <https://trec.nist.gov/pubs/trec16/papers/LEGAL.OVERVIEW16.pdf>; Douglas W. Oard, et al., Overview of the TREC 2008 Legal Track (November 1, 2008), (finding 24% recall), available at <https://apps.dtic.mil/sti/pdfs/ADA512711.pdf>.

*documents were not found ... indicating that it is not just tangentially relevant documents that are being missed by the negotiated Boolean approach.*⁶

Deficiencies of Stacking TAR on Top of Search Terms

14. What happens when TAR is applied to a collection after it has been filtered with search terms? Let's do the math.

15. For example, if the application of search terms has a recall of 24% like the consensus queries in Legal Track 2008⁷ and is followed by a TAR process that achieves a modest recall level of 80%, then the overall recall level would be a plainly inadequate 19.2%, with over 80% of responsive documents missed. If the TAR results were then manually reviewed, assuming a manual review recall of 70%, then the overall level of recall would be even lower: 13.4%, with over 85% of responsive documents missed.

16. Accordingly, when multiple processes are applied in winnowing down a collection, the relevant recall rate is the product of the recall rates of each process, not the recall rate for just one process. As noted by Maura Grossman & Gordon Cormack:

To make matters worse, eDiscovery medicine shows frequently promote the sequential use of two or more information retrieval methods, including Boolean search, TAR, and manual review. ***The net effect of this concoction is to achieve considerably lower recall than any of its constituent parts: when multiple information retrieval methods are used in sequence, overall recall is the product of the recall for each constituent method.*** If keyword culling were to achieve 70 percent recall, the TAR tool were to achieve 80 percent recall, and manual review were to achieve 75 percent recall, the recall of a review effort combining them in sequence would be 70 percent \times 80 percent \times 75 percent = 42 percent. It is possible to quibble with the numbers presented here, but not with the fact that each constituent part is imperfect, and that overall or end-to-end recall is considerably less than the weakest link in the chain.

⁶ Oard, *supra*, at p. 39 (emphasis added).

⁷ Douglas W. Oard, et al., Overview of the TREC 2008 Legal Track 9 (November 1, 2008), available at <https://apps.dtic.mil/sti/pdfs/ADA512711.pdf> (Last accessed Mar. 20, 2024).

When applied sequentially, information retrieval methods—whether Boolean search, TAR, or manual review—will always yield inferior recall. Yet the medicine shows would have us believe that we need to consider only the TAR-tool component in our recall calculations, ignoring relevant documents excluded by keyword culling and/or by post-TAR manual review. This is, at best, an extreme case of moving the goalposts, but more likely, a form of legerdemain.

Maura Grossman & Gordon Cormack, *The eDiscovery Medicine Show*, 18 The Ohio State Technology Law Journal 1, 8 (2021) (footnotes omitted) (emphasis added).⁸

Email Threading Strips Out Substantive Content, Cripples Review Platform Functionality, and Impedes the Examination of Witnesses

17. The practice of suppressing production of all emails which are part of a thread – as a chain of emails and responses – except for a so-called “most inclusive” email which includes fragments of those emails – suppresses substantive original evidence. The fragments of emails included in a “most-inclusive” email omit and obscure substantive evidence visible in the suppressed originals. Thread suppression also strips substantive metadata, rendering commonplace important review processes impossible. Finally, thread suppression makes it impossible to confront a witness with a specific email as it existed by itself at a specific time before it was subsumed in a thread.

18. At the outset, it is important to note that email threading, while often seen as just another form of deduplication, is a far different matter. The specifications and operations of MD-5 and SHA-1 hashing algorithms commonly used for deduplication are public and open, with every step of the process fully and precisely set out, and there are no changeable options or parameters which could vary their output. Every MD-5 or SHA-1 hash algorithm, even when implemented in different code in different languages, applications, programs, or operating

⁸ Available at <https://moritzlaw.osu.edu/sites/default/files/2022-01/THE%20EDISCOVERY%20MEDICINE%20SHOW.pdf>.

systems will always return the same respective results as any other MD-5 or SHA-1 algorithm when applied to identical data.

19. In contrast, email threading programs and algorithms are private and proprietary. There are no agreed public standards or algorithms to which an email threading algorithm must conform. For example, an email threading program may or may not require that all parties to any email in a thread be parties to all the emails in a thread, or that any attachment to any email in a thread must also be attached to every email in a thread.

20. These are critical issues. For example, an email threading algorithm may allow in-thread emails in a thread which contains attachments to be suppressed so long as a so-called “most-inclusive” email, the only email from a thread that would be produced, includes all the attachments that appeared in the thread, even though this would make it impossible to determine when any attachment entered a thread, *which is substantive evidence, critical in establishing who knew what when*.

21. The appearance of in-thread emails in a so-called “most inclusive” email omit and obscure substantive evidence visible in the suppressed originals. For example, attachment names plainly visible at the top of a standalone email get dropped by Outlook when the body of that email is included in a thread.

22. Beyond the suppression of substantive evidence, email thread suppression also disrupts, impedes, and prevents normal litigation preparation.

23. First, the email metadata fields which attorneys will use for searching and review would not include the metadata (From, To, CC, Date Sent, etc.) of the suppressed emails. For example, the only person included in the From metadata field would be the sender of the so-called most-inclusive email. If Plaintiffs’ attorneys search the From field for all emails authored

by a particular custodian, deponent, or witness, the search results will not include suppressed-in-thread emails where that custodian, deponent, or witness is the sender unless, by serendipity, that custodian, deponent or witness also happened to be the sender on the so-called most inclusive email.

24. Similarly, the Date Sent metadata field will be the date sent of the most inclusive email. The Date Sent of the suppressed in-thread emails cannot be searched as they could be if the in-thread suppressed emails were produced separately.

25. Consider the implications in the context of one of the most common and important tasks performed over the entire litigation lifecycle: searching for documents on a particular issue within a particular time window and then sorting them chronologically for review and analysis. *Critically important in-thread emails within that time window would not even be returned in the results* unless the Date Sent of the so-called “most inclusive” email also happened to be within that time window. And even if the Date Sent of the so-called “most inclusive” email did happen to fall with the time window being searched, all of the in-thread mails within it would be mis-sorted. Also, when in-thread emails are produced as separate documents, their content appears right at the top with high visibility for ready comprehension; when included in a thread that content may effectively be buried deep within an extended thread.

26. Moreover, the emails suppressed by email threading have independent significance in their original forms. For example, during depositions or cross-examination at trial, it may be critically important to confront a deponent or witness with an email without disclosing what happened or was said afterward with respect to that email, i.e., to be able to confront the deponent with a specific email as it existed by itself at a specific time before it was

subsumed in a thread. This is impossible when the production of in-thread emails is suppressed because they are not a so-called “most inclusive” email.

27. The stipulated ESI protocol in the *In Re Allergan Biocell Textured Breast Implant Liability Litigation* MDL got it right:

Due to the importance of the metadata in prior or lesser-included emails contained in whole or in part in a most-inclusive email, and in order to facilitate reasonable evidentiary use of emails, production of a most inclusive email thread does not relieve the Producing Party of its obligation to produce responsive prior or lesser-included emails. No document shall be withheld from production on the basis that it is included in a produced more-inclusive email.

In Re Allergan Biocell Textured Breast Implant Liability Litigation, Case 2:19-md-02921, MDL No. 2921 (D.N.J. Sept. 4, 2020), ECF No. 194, Case Management Order No. 15 (Order Regarding Electronically Stored Information and Hard Copy Documents), § B.3.

28. Here, Defendants’ proposal that to “the extent a Receiving Party requests the production of specific lesser-inclusive copies, the Parties will promptly meet and confer and will not unreasonably deny particularized requests” is fatally flawed. Ex. 2, § VI.B. First, it doesn’t correct any of the issues set out above. *A party may not even know that an in-thread email which wasn’t returned in a search with a date range exists.* Second, even if a party was aware of an in-thread email, why should it be forced to disclose its interest in it – i.e., their work product and litigation strategy – in order to obtain it? Third, requiring these requests on a one-off basis would interrupt Plaintiffs’ workflow and ability to search for and work with documents. Fourth, even if Defendants promptly produced a requested document, it would not be available for use in Plaintiffs’ review platform until their vendor loaded it, imposing additional costs and introducing further delays. Fifth, requested documents will *not* be promptly produced; production of a requested document would occur, if at all, only after (i) a meet and confer has been scheduled and (ii) convened, (iii) Defendants agree to produce the document, (iv) their vendor prepares the

record for production and (v) provides it to Plaintiffs' vendor (vi) which would then have to load it into Plaintiffs' review platform.

29. Finally, under Plaintiffs' proposal, Defendants can use email threading for their own internal review should they choose to do so. What they cannot do is obstruct or suppress the production of the individual emails within a thread, and thereby impose crippling burdens on the Plaintiffs' ability to comprehend and use them.

The Necessity of Native Word, PowerPoint, and PDF Documents

30. Plaintiffs' proposal provides for native production of (i) spreadsheets files, (ii) presentation files, (iii) word processing files that contain track changes, comments, hidden text, or embedded file types, (iv) PDF files, and (v) media files be produced in their "native" file format, rather than as black-and-white TIFF images of those files. *See* Ex. 1, § VI.D.⁹

31. Defendants refuse to produce any file types in native format other than (i) spreadsheets, (ii) audio/video/media files, and (iii) unidentified "other files" not conducive to production in image format, meaning that PowerPoint files, Word documents with track changes, and PDFs would all be produced as black-and-white TIFF images. *See* Ex. 2, § VI.D.

32. Word documents, PowerPoint presentations, and PDFs can all contain links. With native documents, a user can hover over a link to reveal the underlying URL, a capability that is lost when those documents are produced as images.

⁹ Plaintiffs propose that the parties meet and confer as to the production format of other document types not specifically listed in their proposed ESI Order. *Id.*

33. This issue may be compounded by some ESI processing tools, which will not extract the URL behind the link and include it in the extracted full text making the URL both invisible and unsearchable¹⁰.

34. In litigation support platforms such as *Relativity*, search hits in Word documents, PowerPoint presentations and text-based PDFs are highlighted in the native document with formatting, color, graphic elements, etc., so that a user can comprehend the full import of the document quickly, navigating as necessary from search hit to search hit within that document at maximum speed and efficiency.

35. But with TIFF+, search hits in Word documents, PowerPoint presentations, and text-based PDFs are not highlighted in the images, but rather in the separate full text (for example, in *Relativity*, there are separate tabbed sections for a document's native file, image, full text, etc.), without any way to jump from a search hit in the full text to the corresponding TIFF image. Instead, users have to estimate the relative position of the search hit in the full text and then estimate what TIFF image page it would be on. If their initial guess is wrong, they then would have to flip through the TIFF images to find the right one, significantly slowing down the review process.

36. What about navigation to the next search hit in a document, a single-click process with native files? With TIFF+, it's not that quick or simple; users can't just click to navigate to the TIFF image page corresponding to the next search hit, but instead would have to:

- a. go back to the separate full text for the document, and then
- b. click to navigate to the next search hit in the separate full text, and then

¹⁰ The text displayed on a link would be captured, so a URL would be captured if and only if it is also the displayed text of a link.

- c. start the estimating all over again to locate the corresponding TIFF image.

Again, this is a major roadblock to efficient review.

37. Tracked changes and comments in Word documents are tough enough to unravel and understand in a native Word document, as anyone who has exchanged draft Word documents knows only too well. That job becomes monumentally harder, slower, and less efficient when the native file, with color, and with inline edit details: *who made a change? when did they make it?* – instantly available by hovering over an edit, is degraded to a black-and-white, or even a color, image.

38. Word documents, PowerPoint presentations, and PDFs are stripped of significant and substantial functionality when degraded to black and white TIFF images. In Word, PowerPoint presentations, and PDF documents, as in other media, color is used, *inter alia*, for impact, to direct the reader's eye, to impart emphasis, and to make a document more readily, i.e., quickly, comprehensible. When a Word document, PowerPoint presentation, or a PDF is degraded to black and white images, these critical factors are lost, and documents accordingly take longer to comprehend.

39. PowerPoints are presentations. When reduced to TIFFs, key elements which give power and impact to a presentation – music, sound effects, animation, etc. – are lost. Graphics and charts, common components in PowerPoint presentations, can become unreadable when degraded to black and white images.

40. PowerPoint presentations can also display graphs, charts, and tables from embedded files, such as an embedded Excel spreadsheet. In native PowerPoint files, a reviewer can double-click on the visible graph, table, chart, etc. and the underlying embedded Excel file will open at the exact spot in the worksheet where the displayed section is, enabling a reviewer to

immediately drill down to reveal the formulas, cross-references, and construction of the displayed section. As Excel files can contain 255 worksheets, each one with up 16,384 columns and 1,048,576 rows, losing this functionality, forcing a user to navigate as best he or she can to find the source section (which can be especially difficult, e.g., when each worksheet has the exact same structure), can impose tremendous burdens.

41. PDFs can also contain multi-media files, which cannot be played from a TIFF.

42. PDF bookmarks function as hyperlinks for immediate direct navigation to particular locations within a PDF. They are often used to create a dynamic table of contents, with immediate navigation to individual chapters or sub-parts. This functionality is also stripped when a PDF is degraded to mere images.

Production and Production Format of Databases and Structured Data

43. My understanding is that pricing data, data that is likely to reside in databases and applications built on databases, will be a core component of this litigation.

44. In my experience, in MDLs where database data is a critical component, ESI protocols provide for negotiation of production and production formats *before* any data is produced. This avoids the downstream delays when unilateral determinations of the data to be produced from a database and the formatting of that data turns out to have flaws which impermissibly restrict the scope and usability of the produced data and have to be redone, creating major expenses and delays.

45. For example, in the VW “Clean Diesel” MDL, the ESI protocol provided:

“Structured or Aggregated data – Databases, Enterprise Systems, and Third Party Cloud Systems.

The Parties will meet and confer to address the identification, production, and production format of any responsive data contained in a database, other structured or aggregated data source, or third party cloud-based storage system. Prior to any

such meet and confer, the Producing Party(ies) will provide sufficient information to enable the Requesting Party(ies) to evaluate the Producing Party's proposed method and format of production. In the event that the Parties cannot reach agreement, the matter may be submitted to the Court for determination."

In Re Volkswagen "Clean Diesel" Marketing, Sales Practices, and Products Liability Litigation, MDL No. 2672, Case 3:15-md-02672 (N.D. Cal. May 4, 2016), ECF No. 1482, Pretrial Order No. 18 (Stipulation and Order Governing the Production of Hard Copy Documents and Electronically Stored Information), § VI.O.

46. Similarly, the ESI Protocol in the 3M Combat Arms MDL provided:

"Databases, Structured, Aggregated or Application Data. The Parties will meet and confer to address the production and production format of any responsive data contained in a database or other structured or aggregated data source or otherwise maintained by an application. The Parties will reasonably cooperate in the exchange of information concerning such databases to facilitate discussions on productions and production format. If the Parties cannot reach agreement, the matter will be decided by the Court or its designee."

In Re 3M Combat Arms Earplug Product Liability Litigation, Case 3:19-md-02885, MDL No. 2885 (N.D. Fla. June 17, 2019), ECF No. 443 Pretrial Order No. 10, Order Governing Production of Documents and Electronically Stored Information at § I.(e).

47. Plaintiffs' proposed ESI Protocol follows this proven, efficient methodology, providing:

"DATABASES AND STRUCTURED, AGGREGATED, OR APPLICATION DATA.

The Parties will meet and confer to address the identification, production, and production format of any relevant or responsive data contained in a database or other Structured Data or aggregated data source or otherwise maintained by an application. The Parties will reasonably cooperate in the exchange of information concerning such databases to facilitate discussions on productions and production format, including available data fields/objects and schema. If the Parties cannot reach agreement, the matter will be decided by the Court or its designee."

Ex. 1, § VIII.

48. Defendants' proposed ESI Protocol turns its back on these proven procedures, cutting a wide swath for unilateral action:

- a. *Databases do not even have to be identified before a party produces data from it.*
- b. The provided meet and confer to address any concerns with the identification, production, and production format of data happens only *after* production, when any miscasting of the data and omissions, etc. will have already been baked in.
- c. It provides that exports and reports shall be produced in a reasonably usable format but that any data produced "in CSV format, tab-delimited text format, Microsoft Excel format, or Microsoft Access format" shall be considered to be in a reasonably usable format, even though, depending on the complexity of the data, some of these formats, particularly when related data from multiple database tables must be produced, could be produced by Defendants in any of these formats, even an unsuitable one like csv or Excel.

See Ex. 2, § VIII.

49. This is a recipe for needless delay, and expense. The databases here will be complex, perhaps including accounting databases where arriving at accurate account and transaction histories may require tracing complicated spider's webs of journal entries, including offsets, credits, etc.

50. Non-accounting databases may have audit tables to record changes and corrections, creating a history which may require delicate and intricate integration with later versions.

51. Given that the parties are likely to have divergent conceptions of scope and relevance, the probability that the Defendants' unilateral decisions will result in inadequate production requiring remediation are unacceptably high.

Native Word Documents, PowerPoint Presentations, and PDFs Are as Easy to Use in Litigation as Native Excel or Audio-Visual Files

52. Native production and use of Excel and audio-visual files is routine and commonplace and has even been accepted by Defendants here in their ESI protocol. *See* Ex. 2, § D.

53. With respect to their post-production treatment in litigation, e.g., for use as exhibits, etc., native Word documents, PowerPoint presentations and PDFs are no different than native Excel or audio-visual files.

54. Native Word, PowerPoint and PDF files can be named with Bates numbers and in the same manner as native Excel or audio-visual files.

55. Single-page Bates-stamped placeholder TIFF images can be produced for native Word, PowerPoint and PDF files in the same manner as native Excel or audio-visual files.

56. Arguments to the contrary are simply illusionary straw men, perhaps full of sound and fury, but ultimately meaning nothing.

Redactions for Relevance Are a Backdoor to Suppression of Evidence

57. Much more than determinations of attorney-client privilege or work product, determinations of relevancy are very much in the eye and subjective judgment of the beholder.

58. Given that the Plaintiffs and Defendants will likely have quite different views on scope and relevancy as applied to Defendants' documents, views which may not be reconciled until much later in this litigation, allowing redactions for relevancy opens the door for abuse because once a Word, PowerPoint or PDF file has been redacted, it no longer has to be produced as a native document.

59. As addressed in ¶¶ 30-42 *supra*, substituting degraded TIFF images for Defendants' native Word, PowerPoint and PDF files would place substantial and significant

burdens on Plaintiffs. Imposition of such burdens should not be permitted on such a subjective and slippery slope as unilateral Defendant determinations of relevancy.

Collecting Modern Attachments as Part of Family Groups is Essential

60. Modern attachments in emails are hyperlinks to documents which are stored in the cloud instead of being physically included¹¹ in an email. They serve the same functional role as a traditional attachment, and appear to be the same to email users:

"Modern attachments

[Cloud repository] OneDrive integrates with Microsoft Outlook to enable easy sharing of OneDrive files that appear just like email attachments."

<https://learn.microsoft.com/en-us/sharepoint/onedrive-overview> (accessed on March 22, 2024).

61. In the Google environment, native Google file types (Google Docs, Google Sheets, etc.) can be attached to Google mail only as modern attachments.

62. Production of a modern attachment as part of a family group with its parent email is every bit as important with modern attachments as it is with traditional attachments.

63. Both Microsoft and Google, the two major current email platforms, provide tools (*Microsoft Purview* for Microsoft and *Google Vault* for Google) for collecting emails along with modern attachments from their respective document repositories.

¹¹ "Physically included" is a simplification with respect to emails resident in within a mailbox, e.g., a .pst or .mbox file, or on a mail server, i.e., anything besides an email message saved as a standalone file, e.g., a .msg file.

64. There are also third-party collection tools, such as Metaspike's *Forensic Email Collector*, which is widely used by governments as well as major law firms and accounting firms.

65. It is my understanding that Defendants have thus far refused to disclose what platforms they are using for email and as document repositories, but regardless of whether Defendants use Microsoft or Google platforms, or both, or other platforms, Defendants should be able to collect email with modern attachments.

I declare under penalty of perjury that the foregoing is true and correct to the best of my knowledge.

DATED: March 22, 2024



Douglas Forrest

Senior Vice President, eDiscovery
Analytics & Strategy
International Litigation Services, Inc.
dforrest@ilsteam.com